www.sciencemag.org/344/6187/998/suppl/DC1

Supplementary Materials for

Labor market returns to an early childhood stimulation intervention in Jamaica

Paul Gertler,* James Heckman, Rodrigo Pinto, Arianna Zanolini, Christel Vermeersch, Susan Walker, Susan M. Chang, Sally Grantham-McGregor

*Corresponding author. E-mail: gertler@haas.berkeley.edu

Published 30 May 2014, *Science* **344**, 998 (2014)
DOI: 10.1126/science.1251178

This PDF file includes:

Contents

A The Jamaican Study

A.1 Intervention and Experimental Design

In 1986-1987, the Jamaican Study enrolled 129 stunted children age 9-24 months that lived in poor disadvantaged neighborhoods of Kingston, Jamaica (*35*). Enrollment was conditioned on stunting because it is an easily and accurately observed indicator of malnutrition that is strongly associated with poor cognitive development (*18*). Stunting was defined using international standards as having a height less than two standard deviations of reference data by age and sex (*45*). The children were stratified by age (above and below 16 months) and sex. Within each stratum, children were sequentially assigned to one of four groups by random assignment. The four groups were (1) psychosocial stimulation (N=32), (2) nutritional supplementation (N=32), (3) both psychosocial stimulation and nutritional supplementation (N=32), and (4) a control group that received neither intervention (N=33). All children were given access to free health care regardless of the group to which they were assigned.

The stimulation intervention (comprising groups 1 and 3) consisted of two years of weekly one-hour play sessions at home with trained community health aides[1] designed to develop child cognitive, language and psychosocial skills. Activities included mediating the environment through labeling, describing objects and actions in the environment, responding to the child's vocalizations and actions, playing educational games, and using picture books and songs that facilitated language acquisition. The first 18 months included Pigetian concepts such as use of a tool and object permanence (*46*). After 18 months concepts such as size, shape, quantity, color and classification based on Palmer (1971) were included. Particular emphasis was placed on the use of praise and giving positive feedback to both the mother and child. Each session's curriculum was adjusted to the child so that activities were at the appropriate level for the child.

A major focus of the weekly visits was on improving the quality of the interaction between mother and child. At every visit the use of homemade toys was demonstrated and the toys were left for the mother and child to use until the next visit when they were replaced with different ones. Mothers were encouraged to continue the activities between visits. The intervention was innovative not only for its focus on structured activities to promote cognitive, language and socio-emotional development but also for its emphasis on supporting the mothers to promote their child's development.

The nutritional intervention (comprising groups 2 and 3) was aimed at compensating for the nutritional deficiencies that may have caused stunting. The nutritional supplements, provided weekly for 18 months, consisted of one kilogram of formula containing 66% of daily-recommended energy (calories), and 100% of daily-recommended protein and micronutrients (see (*47*) for details). In addition, in an attempt to minimize sharing of the formula with other family members, the family also received 0.9 kilograms of cornmeal and skimmed milk powder. Despite this, sharing was common and uptake of the supplement decreased significantly during the intervention (*48*).

Of the 129 study participants, two of the participants dropped out before completion of the two-year program. The remaining 127 participants were surveyed at baseline, resurveyed immediately following the the end of the two-year intervention, and again at ages 7, 11, and 17. Our analysis is based on a re-interview of the sample in 2007-08 when the participants were approximately 22 years old, some 20 years after the original intervention. We obtained 105 interviews at age 22.

A.2 External Comparison Group

For comparison purposes, the study also enrolled a sample of non-stunted children from the same neighborhoods, where non-stunted was defined as having a height for age z-score greater than -1 standard deviations. At baseline,

[1]The aides had completed at least secondary education and training in nutrition and primary heatlh care as part of the government job. They were seconded to the study and received an additional 8 weeks of training in child development, teaching techniques and toy making (*35*).

every fourth stunted child in the study was matched with one non-stunted child who lived nearby and was the same age (plus or minus 3 months) and sex. At age 7, this sample of 32 was supplemented with another 52 children who had been identified in the initial survey as being non-stunted and fulfilled all other inclusion criteria. Members of the non-stunted comparison group did not receive any intervention, but did receive the same free health care as those in the stunted experimental group. From age 7 onwards, this group was surveyed at the same time as the participants in the experiment.

A.3 Previous Studies

The stimulation and the combined stimulation-nutrition arms of the Jamaica Study proved to have a large long-term impact on cognitive development. At the end of the 2-year intervention, the developmental levels of children who received stimulation were significantly above the control group and approached those of the external non-stunted group (*32*). While cognitive benefits decreased somewhat by age 7, significant long-term benefits were sustained through age 22 (*36, 37*). Moreover, stimulation had positive impacts on psychosocial skills, schooling attainment and reduced participation in violent crimes (*36*).

While the stimulation arms had strong and lasting effects, the nutrition-only arm had no long-term effect on any measured outcome (*36, 38*).[2] In addition, there were no statistically significant or quantitatively important differences in effects between the stimulation and stimulation-nutrition arms on any long-term outcome. Hence, we combine the two psychosocial stimulation arms into a single treatment group (N=64) and combine the nutritional supplementation only group with the pure control group into a single control group (N=65).[3] Henceforth we use the term stimulation effects of stunted participants to designate the analysis that compares groups 1 and 3 against groups 2 and 4.

B The New Survey

We resurveyed both stunted (experimental) and non-stunted (comparison) study populations in 2007-08 some 20 years after the original intervention when the participants were approximately 22 years old.[4] We attempted to find all of the study participants regardless of current location and followed migrants to the the US, Canada, and the UK. When we could not find a participant in Jamaica, we contacted relatives for further information to find the participants.

B.1 Stunted Experimental Sample

We were able to find and interview 105 out of the original 127 (83%) stunted participants who completed the program. The stunted sample remained balanced as we only observe significant differences in 3 out of 23 variables (Table S.1). Mothers of children in the treatment group were more likely to be employed and have completed less schooling than mothers of children in the control group, and children in the treatment group had lower weight for height than children in the control group. These imbalances are already present in the full baseline sample of 127,

---

[2]This is in contrast to the Guatemala Study in which nutritional supplementation did affect both long-term health status and earnings ((*44*); (*49*)). Supplementation in Jamaica may have begun too late to have had an impact. The Guatemala study started supplementing children in utero and at birth, before the children became malnourished, while the Jamaican program started at later ages after the children were already malnourished. Other possible reasons for the difference include the fact that the supplement was more intensively shared with other family members in Jamaica and the supplement was a smaller share of the total food budget in Jamaica (*35, 44, 47*).

[3]We formally test the hypotheses that groups 1 and 3 and be pooled, that groups 2 and 4 can pooled, and that supplementation had no impact on earnings in Appendix D.2.

[4]The survey received ethical clearance from the IRB of the University of the West Indies in Kingston, Jamaica.

which suggests that they were the result of sampling variation in the original randomization rather than differential sample attrition. We control for baseline imbalances using Inverse Propensity Weighting (IPW), which re-weights observed data using predicted probabilities of treatment (*39*). The predictions come from a logit model of treatment assignment as a function of the baseline characteristics whose means are significantly different between treatment and control groups.

Twenty-two (17%) of the 127 original participants were not interviewed, of which 10 were not found, 9 died, and 3 of those who were found refused to be interviewed. Of the 13 that were not found or refused to be interviewed, 9 were migrants. Treatment status is not a significant predictor of the overall probability of attrition and the baseline means of none of the 23 individual variables are not significantly different between the group that dropped out and the group that stayed in the sample, even when we stratify by treatment and control (Table S.2). Hence, in terms of measured variables, there appears to be no selective attrition and the remaining sample is representative of the original sample.

We examine the impact of the intervention on densities of log earnings. Figure S.1 presents Epanechnikov kernel density estimates of the treatment and control groups estimated using bandwidths that minimize mean integrated squared error for Gaussian data. The Figure shows the estimated density for the earnings variables associated with first, last and current job.

B.2 Non-Stunted Comparison Sample

We found and interviewed 65 children out of the 84 children originally surveyed with an implied attrition rate of 23%, which is slightly higher than that for the experimental sample. There are, however, significant differences in the baseline characteristics of the attrition and non-attrition groups for 4 out of the 15 variables in the non-stunted sample (SOM Table S.3). Mothers in the attrition group are older, perform better on the Picture Peabody Verbal Test (PPVT), provide more verbal stimulation to their children and live in better houses than mothers who do not attrit. We correct for attrition using IPW to re-weight the observed data using predicted probabilities of attrition (*39*). The predictions come from a logit model of attrition as a function of the baseline characteristics whose means are statistically significantly different between attrited and non-attrited groups.

In order to better understand the external validity of our catch-up analysis we compare the non-stunted group to the general population using data on individuals 21-23 years old living in the greater Kingston area from the 2008 Jamaica Labor Force (JLF) survey that was collected in the same year as the last follow-up. Unfortunately, the labor supply and earnings questions in the JLF and in our survey were asked in different ways, and there was a 50% non-response rate in the JLF to the earnings questions among those who were employed. Only the education variables are directly comparable. By age 22, the non-stunted group attained comparable levels of human capital as those of the same age and living in the Kingston Area interviewed in the Labor Force Survey (SOM Table S.4). The two samples are equally likely to still be in school and achieve the same level of educational attainment in terms of years of schooling and passing national comprehensive matriculation exams. This suggests that the human capital of the non-stunted comparison group is not different from a representative sample of youth in the Kingston area during the study period.

Table S.5 compares education at 22 years old and skills at 18 years old for the non-stunted comparison sample and the stunted sample in the treatment group. The non-stunted comparison sample performs consistently better only in measurements for cognitive skills, but cannot be distinguished from the stunted treated group for all other dimensions.

Figure S.2 presents the Kernel estimates of the earnings densities for the comparison and treatment group. The Figure shows the estimated density for the earnings variables associated with first, last and current job. The next section presents an empirical analysis of baseline variables, attrition and external validity.

B.3 Baseline, Attrition, External Validity and Treatment Effect Tables

This Appendix presents descriptive statistics of baseline variables as well as tests for baseline of the treatment and control stunted sample, selective attrition and external validity.

Table S.1 investigates whether baseline means of the stunted sample are balanced between treatment and control groups. The table reports means of the two groups and the difference in means. The p-values are for two-sided permutation tests of the null hypotheses that the baseline means of the treatment and control groups are equal. We only observe statistically significant differences in 3 out of the 23 variables we examined.

Table S.2 investigates if there is evidence for non-random attrition in the stunted sample. The p-values are for two-sided permutation tests of the null hypotheses that the baseline means of the sample found in the 2008 and the sample not found in 2008 are equal. The first column of the table reports p-values for the full sample and the next two columns report the p-values separately for, respectively, the treatment and control samples. We found no statistically significant differences between the missing and non-missing samples.

Table S.3 investigates if there is evidence for non-random attrition in the non-stunted comparison sample. The p-values for two-sided permutation tests of the null hypotheses that the baseline descriptive statistics for the non-stunted sample found in the 2008 survey (Non-Attrited) and the group lost in the 2008 survey (Attrited) are equal. We observe statistically significant differences in 4 out of the 15 variables we examined.

Table S.4 examines the external validity of non-stunted comparison group. It compares human capital measures from the non-stunted sample collected in 2008 with individuals age 22 and 23 years old living in Kingston Metropolitan Area from from the 2008 Jamaica Labor Force (JLF) survey. The p-values are for for two-sided permutation tests of the null hypotheses that the difference in means between the Jamaican non-stunted sample and the JLF sample is zero.

Table S.6 reports the estimated impacts of treatment on log monthly earnings for the observed sample with imputations for the earnings of missing migrants (9 observations imputed). It displays the analysis of three types of earnings associated with the available data on the participant first job, last job and current job.

Table S.7 examines the catch up effect on Log Earnings between the non-stunted and stunted treatment and control samples. It displays the type of variables examined in Table S.6. Namely, earnings associated with the available data on the participant first job, last job and current job.

C Methodology

We investigate two questions; (1) What is the impact of the stimulation treatment on earnings and (2) Does treatment enable the stunted treatment group to catch-up with the non-stunted comparison group? We estimate the treatment effect on earnings in the experimental stunted sample by linear regression controlling for the variables used in the randomization protocol (age and sex). The catch-up analysis compares the non-stunted comparison group with the stunted treatment group. We estimate the catch-up effects using linear regression also controlling for age and gender.

The small sample size of the Jamaican Study suggests that classical statistical procedures that rely on large sample asymptotic theory to justify the distribution of test statistics may be misleading. We address this problem by using non-parametric permutation tests as implemented in (*21*). Permutation tests are valid in small samples because they are distribution free and do not rely on assumptions about the parametric sampling distribution. The structure of the randomization protocol requires us to permute within the age-sex strata blocks used for the initial randomization. For the treatment effect analysis we expand the number of strata blocks to include the other variables not balanced at baseline.

Increasing the number of blocks of permutations reduces the number of participants that share the same values of the conditioning variables. This may render some permutation blocks invalid as some blocks may contain only treatments or only controls. Effectively, we lose those observations as the treatment status does not vary within this block. To avoid this problem, we apply a parsimonious selection of conditioning covariates in which we only add new covariates besides the ones used in the randomization protocol variables if they significantly explain the outcome of interest. Specifically, we perform a linear regression in which the outcome of interest is explained by the treatment status, the baseline variables used in the randomization protocol and additional variables we ought to examine. We only include the additional variables if we are able to reject the null hypothesis that equates the linear coefficient of the variable being examined to zero. We perform the inference using a double-sided bootstrap p-value and we adopt a significance level of 5%.

Application of the block-permutation test is straightforward for discrete variables, but it requires discretization of continuous variables. Weight-for-height is the only variable that we had to discretize. We chose the largest possible number of divisions that maximize the minimum number of observations in a block. This led us to divide the sample in three categories, those with a z-score higher than 0, those less than 0 but greater than -2, and those less than -2 in the standardized weight for height distribution. We lost no observations for the permutation analysis by following this rule.

The presence of multiple outcomes leads to the potential danger of arbitrarily selecting "statistically significant" outcomes where high values of test statistics arise by chance. Testing each hypothesis one at a time with a fixed significance increases the probability of a type-I error exponentially as the number of outcomes tested grows. We correct for this potential source of bias in inference by performing multiple hypothesis testing based on the Family-Wise Error Rate (FWER), which is the probability of rejecting at least one true null Hypothesis. We use the stepdown algorithm proposed in (*50*), which generates inference exhibiting strong FWER control. Associated with each outcome is a single null hypothesis of no treatment effect. We implement the stepdown procedure for conceptually similar blocks of outcomes.

In addition to the stepdown procedure, we perform multiple hypothesis inference based on a non-parametric aggregator of the outcomes measures. We first transform the data into the relative order of participants across outcomes and then rank each participant within each outcome. We then use the difference in means of participant rank-average as a test statistic. Formally, let $\mathcal{I}$ be the set indexing participants of the Jamaican intervention. Let $D = (D_i; i \in \mathcal{I})$ be the vector of treatment assignments, such that D_i takes value 1 if participant i is assigned to treatment and 0 otherwise. Let $\mathcal{K} = \{1, \ldots, K\}$ be an index set for a selection of outcomes sought to be tested, such that $Y_k = (Y_{i,k}; i \in \mathcal{I})$ denotes the vector of k-th outcome associated with index $k \in \mathcal{K}$. Let Y_k be the dimension of outcome vector Y_k. In this notation, we can compute the rank of the participants within outcome k by:

$$\forall i \in \mathcal{I},\ R_{i,k} = \frac{\sum_{j \in \mathcal{I}} \mathbf{1}[Y_{i,k} \geq Y_{j,k}]}{|Y_k|}.$$

Let the average rank of participant $i \in \mathcal{I}$ across outcomes in $\mathcal{K}$ be:

$$\forall i \in \mathcal{I},\ R_i = \frac{\sum_{k \in \mathcal{K}} R_{i,k}}{|\mathcal{K}|}.$$

The vector of the rank average across outcomes in $\mathcal{K}$ for all participants in $\mathcal{I}$, that is, $R = (R_i; i \in \mathcal{I})$, can be used as a combined measure across outcomes. The associated test statistic comparing treatment and control is the standard difference in means across treatment groups, namely:

$$\Delta R = \frac{\sum_{i \in \mathcal{I}} D_i R_i}{\sum_{i \in \mathcal{I}} D_i} - \frac{\sum_{i \in \mathcal{I}} (1 - D_i) R_i}{\sum_{i \in \mathcal{I}} (1 - D_i)}.$$

We use permutation methods to obtain the sampling distribution.

D Robustness Tests

In this section we report the results of several analyses that examine the extent to which the estimates of treatment on earnings in Table 1 are are robust to a number of assumptions and potential concerns.

We first test the hypothesis that we can pool the stimulation and combined stimulation arms into a single treatment group by estimating the treatment effect on log earnings separately for the pure stimulation intervention and for the combined stimulation/supplemental intervention. We identify the first by comparing the pure stimulation group to the pure control group and the second by comparing the combined stimulation/supplementation to the pure supplementation arm (Tables S.8–S.9). These comparisons ensure that in both cases the only difference between the treatment and comparison groups is the stimulation intervention. The results presented in Tables S.8–S.9 show that the estimated effect sizes of the two approaches are close to one another in almost all cases.[5] Moreover, we cannot reject the hypothesis that the estimated treatment effects using the stimulation group are equal to the estimated treatment effect using the combined stimulation/supplementation group in all cases.

Second, we test the hypothesis that there is no effect of nutritional supplementation on log earnings and that we can pool the supplementation and pure control groups. We estimate the treatment effect on log earnings separately for the pure supplementation intervention and for the combined stimulation/supplemental intervention by comparing the pure supplementation group to the pure control group and the second by comparing the combined stimulation/supplementation to the pure stimulation arm (Tables S.10–S.11). These comparisons ensure that in each case the only difference between the treatment and comparison groups is the supplementation intervention. The results presented in Tables S.10–S.11 show that the estimated effect sizes of the two approaches are in general close to one another and none are statistically significant. In all cases we cannot reject the null hypothesis that the estimated treatment effect using the supplementation group is equal to the estimated treatment effect using the combined stimulation/supplementation group. Finally, we estimate the effect of supplementation on log earnings by comparing the the pooled supplementation and combined supplementation/stimulation groups to the pooled stimulation and control groups (SOM Table S.12) and find that none of the estimated effects are statistically significantly different from zero.

Third, we examine the extent to which the estimates may be affected by censoring in that we only observe the earnings of those employed who are in the labor force. We estimate the effect of treatment on labor force participation using the methods discussed in Section C and find that treatment does not appear to affect overall labor force participation or participation in full or non-temporary jobs (Table S.13). These results are consistent with negligible bias from censoring in the estimated treatment effects on earnings.

Fourth, we examine the concern that the results might be driven by migrants and therefore sensitive to the imputations of earnings for the 9 missing migrants. We re-estimate the effect of treatment on log earnings excluding the migrants (SOM Table S.14) and find that estimated effect sizes remain very close to the original estimates both in terms of magnitude and and statistical significance.

Finally, we assess the extent to which the IPW correction for baseline imbalance affected the estimates by re-estimating the effects of treatment on earnings without the IPW weights (SOM Table S.15). Again, we find the results remain close to the original estimates both in terms of size and significance.

D.1 Empirical Analysis

This appendix reports results of analyses of the robustness of the estimated treatment effects the stimulation on log earnings reported in Table 1 to assumptions and adjustments. Specifically, this appendix presents evidence to support (1) the pooling of the stimulation and combined stimulation/supplementation arms into a single treatment group, (2) that there is that there is no effect of supplementation on earnings, (3) that the estimates are not sensitive

[5]The coefficients are close in 20 out of 22 cases and positive in all cases. Despite the substantially smaller sample sizes than in the pooled model, 10 of the estimated effects are significantly different from zero and the combined rank mean test is statistically significant in 5 out of the 8 cases.

to the treatment of migrants and adjustments to control for imbalances of some characteristics at baseline, and (4) that there are no statistically significant differences by gender.

D.2 Pooling of Stimulation/Supplementation arms

We estimate the treatment effect of stimulation on log earnings separately for the pure stimulation intervention and for the combined stimulation/supplemental arm by comparing the pure stimulation group to the pure control group and the the combined stimulation/supplementation arm to the pure supplementation arm using the methods described in Section C. These comparisons ensure that in each case the only difference between the treatment and comparison groups is the stimulation intervention.

We present these results in Tables S.8–S.9. Results for different earnings indicators are reported in panels A–D. Columns are associated with different job types. Within each panel we first report the results from comparing the stimulation group to the control group and then, just below, the results from comparing the combined stimulation/supplementation group to the supplementation group. Finally, the last row in each panel presents the two-sided p-values for the permutation test of equality of treatment effects estimated with just the pure stimulation group and estimated with the combined stimulation/supplementation group. The estimated effect sizes of the two approaches are close to one another in 20 out of 22 cases and positive in all cases. Despite the small sample sizes, 10 estimates are significantly different from zero and the combined rank mean test is statistically significant in 5 out of the 8 cases. However, the stimulation arm of the intervention shows strong results. It does not matter if we compare stimulation with no treatment or stimulation/suplementation with supplementation only. In both comparisons we examine a group that has stimulation with a group that does not have the stimulation. In both cases it shows strong effects of the stimulation treatment. There is weak evidence that nutrition plus stimulation is more effective for many outcomes than stimulation alone. Finally, in all cases we cannot reject the hypothesis that the estimated treatment effect using the stimulation group is equal to the estimated treatment effect using the combined stimulation/supplementation group.

D.3 The Effect of Nutritional Supplementation on Log-earnings

We estimate the treatment effect of supplementation on log earnings separately for the pure supplementation intervention and for the combined stimulation/supplemental intervention by comparing the pure supplementation group to the pure control group and the combined stimulation/supplementation to the pure stimulation group (Tables S.10–S.11). These comparisons ensure that in each case the only difference between the treatment and comparison groups is the supplementation intervention.

Our analysis is presented in Tables S.10–S.11. Results for different earnings indicators are reported in panels A–D. Columns are associated with different job types. The estimated effect sizes of the two approaches are in general close to one another and slightly negative. None of the 22 estimated effects are statistically significant nor are any of the 8 combined rank mean tests (see SOM section C for a description of this statistic). In all cases we cannot reject the hypothesis that the estimated treatment effect using the supplementation group is equal to the estimated treatment effect using the combined stimulation/supplementation group.

We also estimate the effect of supplementation on log earnings by comparing the the pooled supplementation and combined supplementation/stimulation groups to the pooled stimulation and control groups (Table S.12). None of the estimated effects are significantly different from zero as indicated by the p-values for the individual coefficient estimates and for the combined rank mean tests.

D.4 Adjustments for Migration and Baseline Imbalance

The first 2 tables are used to examine whether the results are sensitive to adjustments for migrants and baseline imbalances in a few characteristics. Specifically, Table S.14 presents the estimated treatment effects excluding all migrants from the sample and Table S.15 presents the estimated treatment effects including migrant but not using

IPW weights to correct for baseline imbalance. In both cases, the estimated coefficients remain very close the estimates in Table 1 in terms of both magnitude and statistical significance.

D.5 Catchup and Migrants

We examine how sensible are our catchup results regarding the subset of stunted migrants. As a robustness check, we re-estimate the catchup model excluding the migrants in Table S.16. We find results are comparable to Table 2 in terms of significance and magnitude.

D.6 Gender Comparison

We also examine the treatment effects of the stimulation arm of the Jamaican intervention separately by gender. We estimated the treatment effects separately by gender (Table S.17) in order to assess whether there are gender differences of the type that have been found in U.S. (*21*). While the estimated effects on earnings are higher for males, tests for equality cannot reject the hypothesis that the impact on earnings is equal for males and females. These results should be considered with caution as the study was neither designed nor powered to assess impact separately by gender. While the estimated effects on earnings are higher for males, tests for equality cannot reject the hypothesis that the impact on earnings is equal for males and females.

E Construction of Earnings Variables

Income data consist of participant reported salaries for each job. The employment survey is job-specific: questions targeted each job separately. Only jobs with positive earnings are considered in our analysis.

In each job, it was asked how many months per year, weeks per months and days per week is the person usually working in that job. Using this data, we classify a job as full time if the participant had at least 20 working days per month. We classify the job as nontemporary and full time if the participant worked full time at least 8 months per year.

We compute total salary for each job by multiplying the total number of months worked by the monthly salary indicated for the job being analyzed. We then divided the overall sum by the number of months worked in total. Average earnings full time is calculated as the sum over all of the full time jobs of monthly earnings multiplied by the number of months spent in each of those jobs , over the sum of all months spent in a full time job.

We also apply a temporal criteria for job classification. Namely, we classify jobs into first, current and last jobs according to employment history of the participant. If two jobs ended in the same year the one yielding the highest earnings was assumed to be the last job. Current job is equal to last job if the person is currently employed. First job is assumed to be the earliest job that ended. If more than one job ended in that year, we define first job as the job with longest duration.

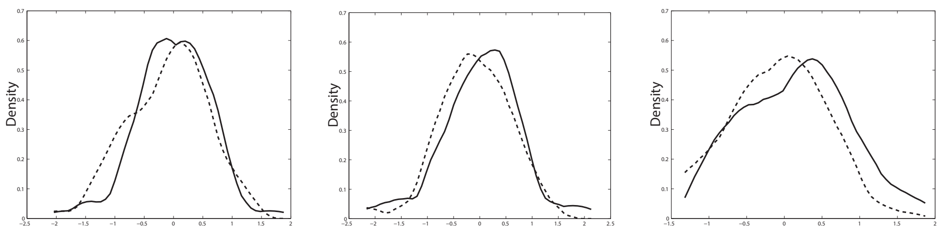
The propensity score is the estimated probability of being in the attrited group at the 22 years old wave. It has been calculated based on three separate models targeting the following samples: (1) stunted treatment group, (2) stunted control group and (3) non-stunted comparison group. Each model was chosen to maximize the Akaike information criteria. The models for the stunted treatment group and for the stunted control group include baseline variables. The majority of the children in the non-stunted comparison group had been added at the 7-years-old follow-up surveys. Thus, the model for the non-stunted comparison group include data from that wave, which is the earliest complete data available.

We use a linear model to impute earning values for missing data regarding migrant workers. Imputation targeted those migrant workers who were lost to follow-up. In the stunted group, these totals of 9 workers. The imputation was obtained through a linear prediction based on selected set of baseline covariates.The small sample size of the data demands a parsimonious covariate selection, which were chosen in order to maximize the Akaike information criteria. Selected covariates are: migrant status, sex and treatment group indicator (for the stunted group) and on migrant status and sex (for non-stunted group).

Figure S.1: Impact of Stimulation Treatment on the Densities of Log Earnings at Age 22

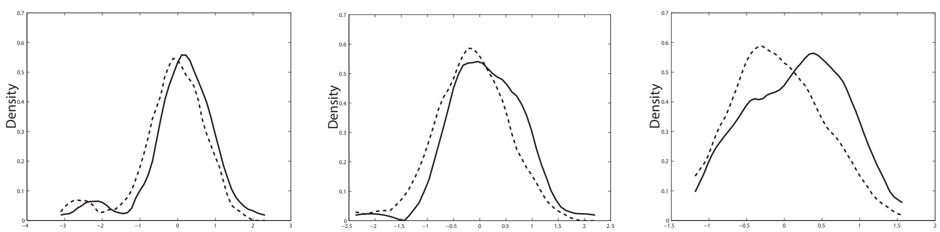
A. Treatment (solid line) and Control (dotted line) Densities for First Job

All Job Types | Full Type Jobs | Non-temporary Jobs



B. Treatment (solid line) and Control (dotted line) Densities for Last Job

All Job Types | Full Type Jobs | Non-temporary Jobs



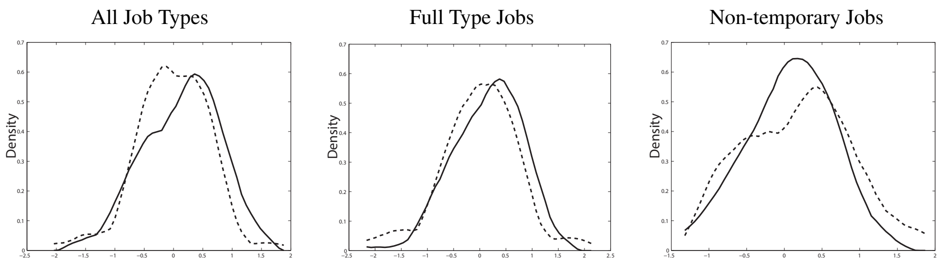
C. Treatment (solid line) and Control (dotted line) Densities for Current Job

All Job Types | Full Type Jobs | Non-temporary Jobs

These figures present the log earnings densities for the treatment and control groups using data where earnings of migrant workers who were lost-to-follow-up were imputed. The control density is the dotted line and the treatment density the solid one. The densities are estimated using Epanechnikov kernels. The treatment densities were estimated with an optimal bandwidth defined as the width that would minimize the mean integrated squared error under the assumption that the data are Gaussian. For comparability purposes, the same bandwidth used was used for the corresponding control group.

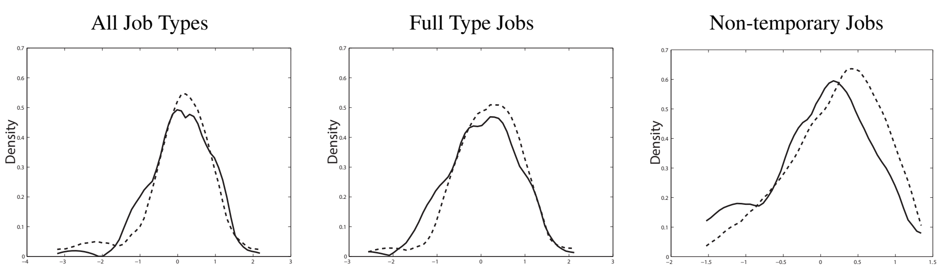
Figure S.2: Catch up of Treatment Group Earnings to Comparison Group Earnings at Age 22

A. Comparison (dotted line) and Treated (solid line) Densities for First Job



B. Comparison (dotted line) and Treated (solid line) Densities for Last Job

C. Comparison (dotted line) and Treated (solid line) Densities for Current Job



These figures present the log earnings densities for the non-stunted comparison and stunted treatment Groups, where where earnings of migrant workers who were lost to follow-up were imputed. The treatment group density is the dotted line and the non-stunted group density the solid one. The densities are estimated using Epanechnikov kernels. The treatment densities were estimated with an optimal bandwidth defined as the width that would minimize the mean integrated squared error under the assumption that the data are Gaussian. For comparability purposes, the same bandwidth used was used for the corresponding control group.

Table S.1: Balance in baseline variables for the observed sample in the 2007/08 follow-up wave

	Sample Size	Control Mean	Treatment Mean	Difference in Means	Single p-value
A. Parental/Household Characteristics					
Mother present	105	0.96	0.94	-0.02	(0.72)
Mother/guardian's age (years)	105	24.35	25.75	1.41	(0.3)
Mother /guardian employed	105	0.15	0.32	0.17	**(0.04)**
Mother/guardian school $\geq$ 9th grade	105	0.21	0.06	-0.15	**(0.01)**
Mother/guardian had any training after school	105	0.23	0.26	0.03	(0.68)
Mother/guardian's PPVT	105	84.88	86.79	1.91	(0.61)
Mothers/guardian's height (cm)	103	159.27	159.33	0.06	(0.96)
Father present	105	0.46	0.45	-0.01	(0.93)
HOME score on enrolment	105	17.13	16.06	-1.08	(0.22)
Access to piped water in the house/yard	105	0.71	0.75	0.04	(0.63)
Household with more than 4 people per room	105	0.35	0.51	0.16	(0.1)
Number of household possessions: at least 3	105	0.25	0.32	0.07	(0.46)
Mother total number of children	105	3.10	3.40	0.30	(0.46)
B. Child Characteristics					
Age (years)	105	1.55	1.55	0.00	(1)
Male	146	0.56	0.53	-0.03	(0.75)
Birth order	105	2.98	3.38	0.40	(0.37)
Birth Weight $<$ 2500 grams	104	0.19	0.25	0.06	(0.49)
Head Circumference (cm)	105	46.24	45.97	-0.27	(0.4)
Daily Calories Consumed	105	1006.78	913.67	-93.11	(0.33)
Daily Protein Consumed (grams)	105	27.05	27.01	-0.04	(0.99)
Griffith Developmental Quotient	105	97.10	99.30	2.21	(0.26)
Height for Age z-Score	105	-2.87	-3.00	-0.13	(0.31)
Weight for Height z-Score	105	-0.87	-1.18	-0.31	**(0.02)**

The table compares means of baseline variables of interest for stunted children in the control group with the ones in the treatment group for the sample observed in 2007-2008. The p-values reported in the last column are for two-sided block permutation tests of the null hypotheses that the difference in means between treatment and control groups are zero. Variable definitions include: PPVT denotes the raw score from Peabody Picture Vocabulary Test (Dunn and Dunn, 1981), HOME denotes the raw score from the HOME environment test (Caldwell, 1967), and Griffith Development Quotient reports the raw score for this test (Griffiths, 1954; Griffiths 1970).

Table S.2: p-Values for Tests of Attrition Bias in the Stunted Sample

	Full Sample	Treatment Group	Control Group
A. Parental/Household Characteristics			
Mother present	(1.00)	(1.00)	(1.00)
Mother/guardian's age (years)	(0.11)	(0.38)	(0.19)
Mother /guardian employed	(0.41)	(0.28)	(1.00)
Mother/guardian school $\geq$ 9th grade	(1.00)	(1.00)	(0.72)
Mother/guardian had any training after school	(0.59)	(0.69)	(0.29)
Mother/guardian's PPVT	(0.55)	(0.60)	(0.75)
Mothers/guardian's height (cm)	(0.86)	(0.55)	(0.51)
Father present	(0.82)	(1.00)	(0.76)
HOME score on enrolment	(0.31)	(0.70)	(1.00)
Access to piped water in the house/yard	(0.28)	(0.45)	(0.48)
Household with more than 4 people per room	(1.00)	(1.00)	(0.51)
Number of household possessions: at least 3	(1.00)	(0.72)	(0.72)
Mother total number of children	(0.41)	(0.67)	(0.51)
B. Child Characteristics			
Age (years)	(0.22)	(0.49)	(0.33)
Male	(0.26)	(0.51)	(0.52)
Birth order	(0.38)	(0.72)	(0.46)
Birth Weight $<$ 2500 grams	(0.23)	(0.16)	(1.00)
Head Circumference (cm)	(0.29)	(0.65)	(0.27)
Daily Calories Consumed	(0.81)	(0.33)	(0.13)
Daily Protein Consumed (grams)	(0.59)	(0.52)	(0.14)
Griffith Developmental Quotient	(0.63)	(0.38)	(0.87)
Height for Age z-Score	(0.37)	(0.77)	(0.29)
Weight for Height z-Score	(0.35)	(0.87)	(0.12)

This table reports the p-values for two-sided permutation tests of the null hypotheses that the difference in baseline means of the sample found in the 2008 and the sample not found in 2008 are equal. The first column reports that results for the full sample and the next two columns report the results separately for, respectively, the treatment and control samples.

Table S.3: Attrition in Non-Stunted group

Variables at the 7-years-old wave	Non-Attrited Group Mean	Attrited Group Mean	Difference in Means	Single p-value
Maternal age	32.38	37.45	5.07	**(0.05)**
Mother present	0.86	0.66	-0.2	(0.13)
Maternal employment	0.66	0.56	-0.1	(0.47)
Maternal education	0.36	0.17	-0.19	(0.10)
Maternal PPVT Score	94.78	84.35	-10.43	**(0.09)**
Home stimulation: books +paper	0.46	0.2	-0.26	(0.30)
Home stimulation: games and trips	0.03	-0.01	-0.04	(0.89)
Home stimulation: verbal stimulation	0.12	-0.3	-0.42	**(0.05)**
Home stimulation: writing material	0.09	-0.06	-0.15	(0.44)
Housing score	8.83	9.56	0.73	**(0.09)**
Child misses school because of money	0.33	0.28	-0.05	(0.77)
Weight for Age z-Score	0.19	0.16	-0.03	(0.88)
Height for Age z-Score	0.81	0.9	0.09	(0.76)
Stanford Binet	82.23	80.74	-1.49	(0.48)
Ravens	13.86	12.84	-1.02	(0.24)

This table reports the baseline descriptive statistics for the sample of non-stunted comparison group member found (Non-Attrited) in the 2008 survey and the group lost (Attrited) in the 2008 survey, using available variables at 7 years old. We used variables at 7 years old because this is the first age where all of the non-stunted children in the final cohort are interviewed (52 non-stunted children were added at 7 years old). The p-values reported in the last 2 column are for two-sided permutation tests of the null hypotheses that the difference in non-attrited and attrited group means are zero.

Table S.4: External Validity of Non-stunted Comparison Group

Comparison with JLFS 2008	JLFS Mean	Comparison Group Mean	Difference in Means	Single p-value
Studying full time	0.09	0.06	-0.03	(0.46)
Highest Grade Completed	10.83	10.87	0.04	(0.76)
Passed at least one CXC exam	0.44	0.36	-0.08	(0.22)
Passed 4 or more CXC exams	0.28	0.32	0.04	(0.33)
Passed at least one CAPE	0.13	0.2	0.07	(0.02)

The table compares the non-stunted comparison group with a sample from the Jamaican Labor Force Survey 2008 (JLFS). The JLFS sample includes individuals of ages 22 and 23 years old living in Kingston Metropolitan Area. The p-values reported in the last column are for two-sided permutation tests of the null hypotheses that the difference in means between the two samples is zero.

Table S.5: Catch Up - Comparison of education and skills for the Non-stunted and stunted treatment samples

	Non-stunted - treatment				
	N	Treatment Mean	β Non-stunted	Single p-Value	SD p-Value
A. Schooling					
Total years of education	114	11.45	-0.11	(0.63)	[0.87]
Any vocational training	118	0.68	-0.08	(0.78)	[0.92]
Any college	118	0.13	0.01	(0.38)	[0.74]
In school	118	0.26	-0.02	(0.57)	[0.86]
In school full time	118	0.19	-0.08	(0.88)	[0.88]
B. Exams					
Passed at least one CXC exam	106	0.33	0.20	(0.01)	[0.03]
Passed 4 or more CXC exams	106	0.22	0.14	(0.05)	[0.09]
Passed at least one CAPE	106	0.09	0.01	(0.37)	[0.37]
C. Cognitive Skills at 18 years old					
WRAT Math	112	31.46	2.44	(**0.01**)	[0.04]
WRAT Reading	112	19.67	4.10	(**0.01**)	[0.03]
Reading-comprehension	112	9.52	1.70	(**0.02**)	[0.04]
PPVT	112	98.85	9.63	(**0.02**)	[0.04]
Verbal Analogies	112	8.96	2.26	(**0.00**)	[0.01]
Ravens Matrices	112	29.60	3.25	(**0.05**)	[0.05]
WAIS IQ	112	72.17	4.97	(**0.01**)	[0.03]
D. Socio-emotional skills at 18 years old					
Oppositional behavior (inverted)	112	-6.02	0.27	(0.34)	[0.50]
Cognitive Problems/Inattention (inv.)	112	-5.65	-0.37	(0.56)	[0.56]
Hyperactivity (inv.)	112	-4.40	1.11	(**0.06**)	[0.15]
Anxiety (inv.)	112	-13.60	0.80	(0.26)	[0.41]
Depression (inv.)	112	-5.81	0.11	(0.43)	[0.43]
Self-esteem (inv.)	112	25.63	0.74	(0.27)	[0.35]

The table presents estimates of the difference in education and skills between the weighted non-stunted comparison group and the stunted cognitive stimulation group. Our p-values are for one-sided block permutation tests of the null hypothesis of complete catch-up on each outcome (single p-value, in parentheses) and accounting for multiple hypotheses (stepdown p-values, in brackets). Permutation blocks are based on gender only, but do not control for differences in baseline values because the aim is to test for catch-up despite the initial disadvantage. Treatment mean is the mean of the stunted group receiving cognitive stimulation. The β is the coefficient for being in the non stunted group after reweighting the data with IPW to correct for attrition. Exams have a smaller sample size because they are only considered for those who did not migrate or migrated after turning 18 years old. Skills have been inverted for oppositional behavior, cognitive problems, anxiety, depression so that they are positively coded.

Table S.6: Treatment Effect on Log Earnings at Age 22

Job Type	All Job Types	Full Time Job	Non-Temporary Job	Combined (Rank Mean)
A. First Job				
Treatment Effect	0.17	0.21	**0.43**	0.09
Single p-value	(0.21)	(0.14)	**(0.04)**	(0.14)
Stepdown p-value	[0.21]	[0.19]	**[0.07]**	–
Sample Size	106	106	83	109
B. Last Job				
Treatment Effect	0.13	**0.27**	**0.41**	**0.08**
Single p-value	(0.23)	**(0.02)**	**(0.01)**	**(0.07)**
Stepdown p-value	[0.23]	**[0.04]**	**[0.02]**	–
Sample Size	110	107	83	110
C. Current Job				
Treatment Effect	0.19	**0.31**	–	**0.09**
Single p-value	(0.21)	**(0.05)**	–	**(0.10)**
Stepdown p-value	[0.21]	**[0.08]**	–	–
Sample Size	81	71	–	81

This table reports the estimated impacts of treatment on log monthly earnings for the observed sample with imputations for the earnings of missing migrants (9 observations imputed). Estimates are not reported for the current job for non-temporary workers because the non-missing sample size is less than 40% of the total sample. The treatment effects are interpreted as the differences in the means of log earnings between the stunted treatment and stunted control groups conditional on baseline values of child age, gender, weight-for-height z-score, maternal employment, and maternal education. Our p-values are for one-sided block permutation tests of the null hypothesis of no treatment effect (single p-value, in parenthesis) and multiple hypotheses (stepdown p-value, in brackets) of no treatment. Permutation blocks are based on the conditioning variables used in the treatment effect regressions. The last column uses a combined statistic that summarizes the participant's outcomes. We perform a a single-hypothesis inference using the average rank across variables as a test statistic. See SOM section C for details.

Table S.7: Catch Up - Comparison of Log Earning of the non-stunted and stunted treatment and control samples

	Panel (I) Non-stunted - treatment				Panel (II) Non-stunted - control			
Job Type	All Job Types	Full Time Job	Non-Temporary Job	Combined (Rank Mean)	All Job Types	Full Time Job	Non-Temporary Job	Combined (Rank Mean)
A. First Job								
Treatment Effect	0.11	0.14	-0.16	0.04	0.21	0.26	0.19	0.09
Single p-value	(0.17)	(0.14)	(0.85)	(0.24)	**(0.05)**	**(0.02)**	**(0.09)**	**(0.03)**
Stepdown p-value	[0.27]	[0.25]	[0.85]	–	**[0.08]**	**[0.03]**	**[0.09]**	–
Sample Size	117	114	97	119	119	120	102	122
B. Last Job								
Treatment Effect	0.09	-0.06	-0.26	-0.02	0.23	0.21	0.10	0.06
Single p-value	(0.30)	(0.67)	(0.95)	(0.65)	**(0.06)**	**(0.06)**	(0.26)	**(0.10)**
Stepdown p-value	[0.46]	[0.75]	[0.95]	–	**[0.10]**	**[0.09]**	[0.26]	–
Sample Size	120	116	96	120	122	121	101	122
C. Current Job								
Treatment Effect	-0.02	-0.21	–	-0.04	0.13	0.09	–	0.05
Single p-value	(0.54)	(0.88)	–	(0.73)	(0.23)	(0.30)	–	(0.21)
Stepdown p-value	[0.51]	[0.85]	–	–	[0.31]	[0.30]	–	–
Sample Size	97	87	–	97	92	86	–	92

The table presents estimates of the difference in the means of log earnings between respectively (I) the weighted non-stunted comparison group and the stunted cognitive stimulation group and (II) the weighted non-stunted comparison group and the stunted control group. Our p-values are for one-sided block permutation tests of the null hypothesis of complete catch-up on each outcome (single p-value, in parentheses) and accounting for multiple hypotheses (stepdown p-values, in brackets). Permutation blocks are based on gender only, but do not control for differences in baseline values because the aim is to test for catch up despite the initial disadvantage. The "combined" column uses the average rank across participants as a test statistic. See Section C of SOM for details.

Table S.8: Treatment Effects of the Stimulation Arms on Log Earnings at Age 22 (Part I)

Job Type	All Job Types	Full Time Job	Non-Temporary Job	Combined (Rank Mean)
A. Average	Stimulation v.s. No-treatment			
Treatment Effect	0.36	0.30	0.41	0.12
Single p-value	**(0.03)**	**(0.06)**	**(0.05)**	**(0.06)**
Stepdown p-value	**[0.06]**	**[0.06]**	**[0.07]**	–
Sample Size	54	52	39	54
	Stimulation/Supplementation v.s. Supplementation Only			
Treatment Effect	0.30	0.22	0.37	0.10
Single p-value	**(0.06)**	(0.11)	**(0.04)**	**(0.08)**
Stepdown p-value	**[0.09]**	[0.11]	**[0.07]**	–
Sample Size	57	54	44	57
Test for Equality of Treatment Effects	0.88	0.78	0.88	
B. First Job	Stimulation v.s. No-treatment			
Treatment Effect	0.11	0.23	0.57	0.10
Single p-value	(0.25)	(0.12)	**(0.02)**	**(0.08)**
Stepdown p-value	[0.25]	[0.16]	**[0.04]**	–
Sample Size	54	53	39	54
	Stimulation/Supplementation v.s. Supplementation Only			
Treatment Effect	0.11	0.08	0.23	0.04
Single p-value	(0.31)	(0.37)	(0.15)	(0.33)
Stepdown p-value	[0.39]	[0.37]	[0.26]	–
Sample Size	53	54	45	56
Test for Equality of Treatment Effects	0.89	0.56	0.27	

This table reports the estimated impacts of treatment on log monthly earnings for the observed sample with imputations for the earnings of missing migrants (9 observations imputed). Outcomes are grouped into block according to its meaning. Block (A) presents the Average Lifetime Earning. Block (B) presents the Log Earnings of the First Job. Each block presents two sets of results. The first one compares the Stimulation Arm with the No-treatment Arm of the Jamaican Intervention. The second set compares the Stimulation and Supplementation Arm with the Supplementation Only Arm of the Jamaican Intervention. Columns report treatment effect estimates for the following job types: All workers, Full Time Workers, and Full Time Non-Temporary workers. The fourth column uses a combined statistic that summarizes the participant's outcomes. We compute the average rank of each participant across the outcomes we examine. See SOM section C for details. The first line of each outcome block present the treatment effect estimate conditional on gender. The second line presents the one-sided p-value for the block permutation test of the null hypothesis of no treatment effect (single p-value, in parenthesis). The third line presents the stepdown p-value (in brackets) associated with the multiple hypotheses testing of no treatment effects. Permutation blocks are based on the conditioning variables used in the treatment effect regressions. Last line of each outcome block presents the double-sided permutation p-value for the test of equality of treatment effects between the groups defined in each panel.

Table S.9: Treatment Effects of the Stimulation Arms on Log Earnings at Age 22 (Part II)

Job Type	All Job Types	Full Time Job	Non-Temporary Job	Combined (Rank Mean)
C. Last Job	Stimulation v.s. No-treatment			
Treatment Effect	0.22	0.30	0.38	0.10
Single p-value	(0.22)	**(0.06)**	**(0.06)**	(0.11)
Stepdown p-value	[0.22]	**[0.09]**	[0.12]	–
Sample Size	54	53	39	54
	Stimulation/Supplementation v.s. Supplementation Only			
Treatment Effect	0.26	0.38	0.46	0.11
Single p-value	(0.12)	**(0.03)**	**(0.01**	**(0.06**
Stepdown p-value	[0.12]	**[0.04]**	**[0.02**	–
Sample Size	58	55	45	58
Test for Equality of Treatment Effects	0.90	0.73	0.77	
D. Current Job	Stimulation v.s. No-treatment			
Treatment Effect	0.15	0.37	–	0.03
Single p-value	(0.32)	(0.11)	–	(0.38)
Stepdown p-value	[0.32]	[0.15]	–	–
Sample Size	37	32	–	37
	Stimulation/Supplementation v.s. Supplementation Only			
Treatment Effect	0.20	0.37	–	0.14
Single p-value	(0.22)	**(0.05)**	–	**(0.05)**
Stepdown p-value	[0.22]	**[0.09]**	–	–
Sample Size	46	39	–	46
Test for Equality of Treatment Effects	0.91	0.96		

This table reports the estimated impacts of treatment on log monthly earnings for the observed sample with imputations for the earnings of missing migrants (9 observations imputed). Outcomes are grouped into block according to its meaning. Block (C) presents the Log Earnings of the Last Job. Block (D) presents the Log Earnings of the Current Job. Each block presents two sets of results. The first one compares the Stimulation Arm with the No-treatment Arm of the Jamaican Intervention. The second set compares the Stimulation and Supplementation Arm with the Supplementation Only Arm of the Jamaican Intervention. Columns report treatment effect estimates for the following job types: All workers, Full Time Workers, and Full Time Non-Temporary workers. The fourth column uses a combined statistic that summarizes the participant's outcomes. We compute the average rank of each participant across the outcomes we examine. See SOM section C for details. The first line of each outcome block present the treatment effect estimate conditional on gender. Estimates are not reported for the current job for non-temporary workers because the non-missing sample size is less than 40% of the total sample. The second line presents the one-sided p-value for the block permutation test of the null hypothesis of no treatment effect (single p-value, in parenthesis). The third line presents the stepdown p-value (in brackets) associated with the multiple hypotheses testing of no treatment effects. Permutation blocks are based on the conditioning variables used in the treatment effect regressions. Last line of each outcome block presents the double-sided permutation p-value for the test of equality of treatment effects between the groups defined in each panel.

Table S.10: Treatment Effects of the Supplementation Arms on Log Earnings at Age 22 (Part I)

Job Type	All Job Types	Full Time Job	Non-Temporary Job	Combined (Rank Mean)
A. Average	Supplementation v.s. No-treatment			
Treatment Effect	-0.12	-0.11	-0.05	-0.06
Single p-value	0.73	0.71	0.59	0.78
Stepdown p-value	0.73	0.78	0.70	–
Sample Size	55	54	43	55
	Stimulation/Supplementation v.s. Stimulation Only			
Treatment Effect	-0.17	-0.18	-0.10	-0.07
Single p-value	0.83	0.84	0.67	0.81
Stepdown p-value	0.90	0.84	0.82	–
Sample Size	56	52	40	56
Test for Equality of Treatment Effects	0.64	0.89	0.78	
B. First Job	Supplementation v.s. No-treatment			
Treatment Effect	-0.11	-0.11	0.07	-0.04
Single p-value	0.71	0.70	0.37	0.68
Stepdown p-value	0.71	0.77	0.50	–
Sample Size	54	56	44	56
	Stimulation/Supplementation v.s. Stimulation Only			
Treatment Effect	-0.11	-0.26	-0.27	-0.10
Single p-value	0.79	0.91	0.89	0.94
Stepdown p-value	0.93	0.91	0.96	–
Sample Size	53	51	40	54
Test for Equality of Treatment Effects	0.69	0.97	0.61	

This table reports the estimated impacts of treatment on log monthly earnings for the observed sample with imputations for the earnings of missing migrants (9 observations imputed). Outcomes are grouped into block according to its meaning. Block (A) presents the Average Lifetime Earning. Block (B) presents the Log Earnings of the First Job. Each block presents two sets of results. The first one compares the Stimulation Arm with the No-treatment Arm of the Jamaican Intervention. The second set compares the Stimulation and Supplementation Arm with the Supplementation Only Arm of the Jamaican Intervention. Columns report treatment effect estimates for the following job types: All workers, Full Time Workers, and Full Time Non-Temporary workers. The fourth column uses a combined statistic that summarizes the participant's outcomes. We compute the average rank of each participant across the outcomes we examine. See SOM section C for details. The first line of each outcome block present the treatment effect estimate conditional on gender. The second line presents the one-sided p-value for the block permutation test of the null hypothesis of no treatment effect (single p-value, in parenthesis). The third line presents the stepdown p-value (in brackets) associated with the multiple hypotheses testing of no treatment effects. Permutation blocks are based on the conditioning variables used in the treatment effect regressions. Last line of each outcome block presents the double-sided permutation p-value for the test of equality of treatment effects between the groups defined in each panel.

Table S.11: Treatment Effects of the Supplementation Arms on Log Earnings at Age 22 (Part II)

Job Type	All Job Types	Full Time Job	Non-Temporary Job	Combined (Rank Mean)
C. Last Job	Supplementation v.s. No-treatment			
Treatment Effect	-0.12	-0.18	-0.07	-0.06
Single p-value	0.68	0.82	0.64	0.78
Stepdown p-value	0.77	0.82	0.81	–
Sample Size	56	56	44	56
	Stimulation/Supplementation v.s. Stimulation Only			
Treatment Effect	-0.08	-0.09	0.01	-0.04
Single p-value	0.61	0.68	0.49	0.67
Stepdown p-value	0.71	0.68	0.67	–
Sample Size	56	52	40	56
Test for Equality of Treatment Effects	0.46	0.35	0.46	
D. Current Job	Supplementation v.s. No-treatment			
Treatment Effect	-0.20	-0.10	–	-0.13
Single p-value	0.78	0.72	–	0.92
Stepdown p-value	0.78	0.81	–	–
Sample Size	38	35	–	38
	Stimulation/Supplementation v.s. Stimulation Only			
Treatment Effect	-0.17	-0.12	–	-0.03
Single p-value	0.71	0.69	–	0.64
Stepdown p-value	0.71	0.81	–	–
Sample Size	45	36	–	45
Test for Equality of Treatment Effects	0.68	0.71		

This table reports the estimated impacts of treatment on log monthly earnings for the observed sample with imputations for the earnings of missing migrants (9 observations imputed). Outcomes are grouped into block according to its meaning. Block (C) presents the Log Earnings of the Last Job. Block (D) presents the Log Earnings of the Current Job. Each block presents two sets of results. The first one compares the Stimulation Arm with the No-treatment Arm of the Jamaican Intervention. The second set compares the Stimulation and Supplementation Arm with the Supplementation Only Arm of the Jamaican Intervention. Columns report treatment effect estimates for the following job types: All workers, Full Time Workers, and Full Time Non-Temporary workers. The fourth column uses a combined statistic that summarizes the participant's outcomes. We compute the average rank of each participant across the outcomes we examine. See SOM section C for details. The first line of each outcome block present the treatment effect estimate conditional on gender. Estimates are not reported for the current job for non-temporary workers because the non-missing sample size is less than 40% of the total sample. The second line presents the one-sided p-value for the block permutation test of the null hypothesis of no treatment effect (single p-value, in parenthesis). The third line presents the stepdown p-value (in brackets) associated with the multiple hypotheses testing of no treatment effects. Permutation blocks are based on the conditioning variables used in the treatment effect regressions. Last line of each outcome block presents the double-sided permutation p-value for the test of equality of treatment effects between the groups defined in each panel.

Table S.12: Treatment Effect on Log Earnings at Age 22 (Supplementation Treatment Effect)

Job Type	All Job Types	Full Time Job	Non-Temporary Job	Combined (Rank Mean)
A. Average				
Treatment Effect	-0.18	-0.18	-0.11	-0.08
Single p-value	(0.84)	(0.86)	(0.65)	(0.84)
Stepdown p-value	[0.89]	[0.86]	[0.77]	–
Sample Size	109	105	82	109
B. First Job				
Treatment Effect	-0.13	-0.19	-0.11	-0.08
Single p-value	(0.79)	(0.90)	(0.87)	(0.93)
Stepdown p-value	[0.85]	[0.90]	[0.87]	–
Sample Size	106	106	83	109
C. Last Job				
Treatment Effect	-0.11	-0.19	-0.09	-0.06
Single p-value	(0.71)	(0.89)	(0.59)	(0.83)
Stepdown p-value	[0.81]	[0.89]	[0.78]	–
Sample Size	110	107	83	110
D. Current Job				
Treatment Effect	-0.27	-0.25	–	-0.11
Single p-value	(0.88)	(0.87)	–	(0.93)
Stepdown p-value	[0.88]	[0.93]	–	–
Sample Size	81	71	–	81

This table reports the estimated impacts of treatment on log monthly earnings for the observed sample with imputations for the earnings of missing migrants (9 observations imputed). We examine the Supplementation Treatment Effects. For treatment group, we combine the supplementation only and the stimulation and supplementation arms of the intervention. For control group, we combine the no-treatment and the stimulation only arms of the Jamaican intervention. We use inverse probability weighting to control for attrition and baseline imbalance correction. Estimates are not reported for the current job for non-temporary workers because the non-missing sample size is less than 40% of the total sample. The treatment effects are interpreted as the differences in the means of log earnings between the stunted treatment and stunted control groups conditional on baseline values of child age, gender, weight-for-height z-score, maternal employment, and maternal education. Our p-values are for one-sided block permutation tests of the null hypothesis of no treatment effect (single p-value, in parenthesis) and multiple hypotheses (stepdown p-value, in brackets) of no treatment. Permutation blocks are based on the conditioning variables used in the treatment effect regressions. The last column uses a combined statistic that summarizes the participant's outcomes. We perform a a single-hypothesis inference using the average rank across variables as a test statistic. See SOM section C for details.

Table S.13: Impact of Stimulation Treatment on Employment and Labor Force Participation at Age 22

	Currently Employed	Full time Job	Non-Temporary Job	Combined (Rank Mean)
Treatment Effect	0.12	0.07	0.09	0.03
Single p-value	**(0.09)**	(0.19)	(0.11)	(0.13)
Stepdown p-value	[0.25]	[0.58]	[0.27]	–
Sample Size	102	103	103	103

The table presents the estimated impact of treatment on labor market outcomes for the stunted experimental sample. The treatment effects are interpreted as the differences in the means of employment outcomes between the stunted treatment and stunted control groups conditional on baseline values of child age, gender, weight-for-height z-score, plus maternal employment and maternal education whenever their contribution in explaining the outcome was statistically significant at the 0.1 level. Our p-values are for one-sided block permutation tests of the null hypothesis of no treatment effect (single p-value, in parenthesis) and multiple hypotheses (stepdown p-value, in brackets) of no treatment. Permutation blocks are based on the conditioning variables used in the treatment effect regressions. The last column uses a combined statistic that summarizes the participant's outcomes. We perform a a single-hypothesis inference using the average rank across variables as a test statistic. See SOM section C for details.

Table S.14: Treatment Effect on Log Earnings at Age 22 Excluding Migrants for the Stunted Experimental Sample

Job Type	All Job Types	Full Time Job	Non-Temporary Job	Combined (Rank Mean)
A. Average				
Treatment Effect	**0.28**	**0.20**	**0.36**	**0.09**
Single p-value	**(0.01)**	**(0.03)**	**(0.02)**	**(0.04)**
Stepdown p-value	**[0.03]**	**[0.03]**	**[0.03]**	–
Sample Size	88	85	62	88
B. First Job				
Treatment Effect	**0.25**	**0.26**	**0.46**	**0.13**
Single p-value	**(0.04)**	**(0.04)**	**(0.01)**	**(0.02)**
Stepdown p-value	**[0.05]**	**[0.04]**	**[0.02]**	–
Sample Size	85	86	63	88
C. Last Job				
Treatment Effect	**0.04**	**0.25**	**0.38**	**0.06**
Single p-value	(0.35)	**(0.03)**	**(0.01)**	(0.11)
Stepdown p-value	[0.35]	**[0.05]**	**[0.04]**	–
Sample Size	89	87	63	89
D. Current Job				
Treatment Effect	0.12	**0.30**	–	**0.07**
Single p-value	(0.29)	**(0.05)**	–	(0.16)
Stepdown p-value	[0.29]	**[0.08]**	–	–
Sample Size	83	71	–	83

The table reports the results of the analyses reported in Table 1 excluding all migrants from the sample. We correct for attrition and imbalance of pre-program variables using the standard method of inverse probability weighting. We replicate the analysis of Table 1 which estimates the impact of treatment on the stunted experimental sample. The last column uses a combined statistic that summarizes the participant's outcomes. We perform a a single-hypothesis inference using the average rank across variables as a test statistic. See SOM section C for details.

Table S.15: Treatment Effect on Log Earnings (No IPW Correction)

Job Type	All Job Types	Full Time Job	Non-Temporary Job	Combined (Rank Mean)
A. Average				
Treatment Effect	**0.34**	**0.26**	**0.42**	**0.11**
Single p-value	**(0.01)**	**(0.02)**	**(0.01)**	**(0.02)**
Stepdown p-value	**[0.01]**	**[0.02]**	**[0.01]**	–
Sample Size	111	106	83	111
B. First Job				
Treatment Effect	0.19	**0.25**	**0.45**	**0.10**
Single p-value	(0.17)	**(0.08)**	**(0.02)**	**(0.09)**
Stepdown p-value	[0.17]	[0.11]	**[0.04]**	–
Sample Size	107	107	84	110
C. Last Job				
Treatment Effect	**0.23**	**0.33**	**0.45**	**0.10**
Single p-value	**(0.08)**	**(0.01)**	**(0.00)**	**(0.02)**
Stepdown p-value	**[0.08]**	**[0.01]**	**[0.01]**	–
Sample Size	112	108	84	112
D. Current Job				
Treatment Effect	0.21	**0.37**	–	**0.11**
Single p-value	(0.14)	**(0.02)**	–	**(0.05)**
Stepdown p-value	[0.14]	**[0.03]**	–	–
Sample Size	83	71	–	83

This table reports the estimated impacts of treatment on log monthly earnings for the observed sample with imputations for the earnings of missing migrants (9 observations imputed). Estimates are not reported for the current job for non-temporary workers because the non-missing sample size is less than 40% of the total sample. The treatment effects are estimated by linear regression and are interpreted as the differences in the means of log earnings between the stunted treatment and stunted control groups conditional on baseline values of child age, gender, weight-for-height z-score, maternal employment, and maternal education. Our p-values are for one-sided block permutation tests of the null hypothesis of no treatment effect (single p-value, in parenthesis) and multiple hypotheses (stepdown p-value, in brackets) of no treatment. Permutation blocks are based on the conditioning variables used in the treatment effect regressions. The last column uses a combined statistic that summarizes the participant's outcomes. We perform a a single-hypothesis inference using the average rank across variables as a test statistic. See SOM section C for details.

Table S.16: Catch Up - Comparison of Log Earning Excluding Migrants
for the Non-stunted and stunted treatment and control samples

	Panel (I) Non-stunted - treatment				Panel (II) Non-stunted - control			
Job Type	All Job Types	Full Time Job	Non-Temporary Job	Combined (Rank Mean)	All Job Types	Full Time Job	Non-Temporary Job	Combined (Rank Mean)
A. Average								
Treatment Effect	0.03	0.02	-0.09	0.04	0.24	0.17	0.16	0.09
Single p-value	(0.40)	(0.44)	(0.73)	(0.28)	**(0.04)**	**(0.09)**	(0.15)	**(0.02)**
Stepdown p-value	[0.53]	[0.50]	[0.73]	–	**[0.06]**	[0.11]	[0.15]	–
Sample Size	97	94	75	97	103	101	83	103
B. First Job								
Treatment Effect	0.11	0.17	-0.01	0.07	0.23	0.30	0.24	0.11
Single p-value	(0.20)	**(0.09)**	(0.52)	(0.11)	**(0.04)**	**(0.01)**	**(0.02)**	**(0.01)**
Stepdown p-value	[0.29]	[0.17]	[0.52]	–	**[0.02]**	**[0.04]**	**[0.04]**	–
Sample Size	94	92	75	96	101	102	84	104
C. Last Job								
Treatment Effect	0.24	0.01	-0.13	0.03	0.27	0.22	0.13	0.08
Single p-value	(0.10)	(0.48)	(0.78)	(0.32)	**(0.04)**	**(0.05)**	(0.20)	**(0.07)**
Stepdown p-value	[0.19]	[0.57]	[0.78]	–	**[0.08]**	**[0.08]**	[0.20]	–
Sample Size	97	94	74	97	104	103	83	104
D. Current Job								
Treatment Effect	0.12	-0.14	–	0.01	0.16	0.10	–	0.05
Single p-value	(0.28)	(0.76)	–	(0.44)	(0.20)	(0.28)	–	(0.23)
Stepdown p-value	[0.37]	[0.76]	–	–	[0.28]	[0.28]	–	–
Sample Size	75	67	–	75	77	71	–	77

The table reports the results of the analyses reported in Table 2 excluding all migrants from the sample. We correct for attrition and imbalance of pre-program variables using the standard method of inverse probability weighting. We replicate the analysis of Table 2 which estimates the impact of treatment on the non-stunted experimental sample. Definitions are as given in Table 2.

Table S.17: Treatment Effect and Gender Differences

	All job Types			Full Time jobs			Non-Temporary jobs		
Job Type	Female	Male	p-Value	Female	Male	p-Value	Female	Male	p-Value
A. Average									
Treatment Effect	0.29	0.44	0.17	0.34	0.24	-0.05	0.49	0.37	-0.09
Single p-value	(0.21)	(0.01)	(0.53)	(0.11)	(0.09)	(0.85)	(0.04)	(0.05)	(0.79)
Stepdown p-value	[0.21]	[0.02]		[0.15]	[0.09]		[0.09]	[0.06]	
Sample Size	50	61		47	59		32	51	
B. First Job									
Treatment Effect	0.17	0.08	-0.06	0.28	0.09	-0.10	0.45	0.37	-0.04
Single p-value	(0.4)	(0.33)	(0.81)	(0.13)	(0.33)	(0.73)	(0.04)	(0.05)	(0.92)
Stepdown p-value	[0.4]	[0.33]		[0.18]	[0.41]		[0.04]	[0.11]	
Sample Size	47	60		47	60		32	52	
C. Last Job									
Treatment Effect	0.18	0.29	0.06	0.46	0.26	-0.17	0.53	0.40	-0.08
Single p-value	(0.18)	(0.15)	(0.84)	(0.01)	(0.12)	(0.55)	(0.01)	(0.04)	(0.81)
Stepdown p-value	[0.18]	[0.15]		[0.02]	[0.17]		[0.02]	[0.07]	
Sample Size	50	62		47	61		32	52	
D. Current Job									
Treatment Effect	0.27	0.16	-0.17	0.66	0.16	-0.55	–	–	
Single p-value	(0.14)	(0.30)	(0.68)	(0.00)	(0.27)	(0.12)	–	–	
Stepdown p-value	[0.14]	[0.30]		[0.00]	[0.36]		–	–	
Sample Size	37	46		28	43				

The table compares the log monthly earnings for stunted cognitive stimulation group and stunted control group stratifying the sample by gender and testing for equality of coefficients between males and females. Treatment effects are reported separately for females (first column in each block) and for males (second column in each block), and p-values correspond to the permuted t-statistic for the hypothesis that the coefficients are the same for males and females. Row blocks analyze: (A) Average Lifetime Earning over all jobs, (B) First Job, (C) Last Job, and (D) Current job . For each type of job, results are reported for the following types of workers as indicated by the column blocks: All workers, Full Time Workers, and Full Time Non-Temporary workers. Estimates are missing whenever the sample reached less than 40% of the total sample. The treatment effects are estimated by linear regression and are interpreted as the differences in the means of log earnings between the stunted treatment and stunted control groups conditional on baseline values of child age, sex, weight-for-height z-score dummies, plus maternal employment and maternal education whenever their contribution in explaining the outcome was statistically significant at the 0.1 level. Our p-values are for one-sided block permutation tests of the null hypothesis of no treatment effect (single p-value, in parentheses) and multiple hypotheses (stepdown p-value, in brackets) of no treatment. Permutation blocks are based on the conditioning variables used in the treatment effect regressions.

References and Notes

1. P. R. Huttenlocher, Synaptic density in human frontal cortex - developmental changes and effects of aging. *Brain Res.* **163**, 195–205 (1979). doi:10.1016/0006-8993(79)90349-4 Medline

2. P. R. Huttenlocher, *Neural Plasticity: The Effects of Environment on the Development of the Cerebral Cortex* (Harvard Univ. Press, Cambridge, MA, 2002).

3. R. A. Thompson, C. A. Nelson, Developmental science and the media. Early brain development. *Am. Psychol.* **56**, 5–15 (2001). Medline doi:10.1037/0003-066X.56.1.5

4. E. I. Knudsen, J. J. Heckman, J. L. Cameron, J. P. Shonkoff, Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10155–10162 (2006). Medline doi:10.1073/pnas.0600888103

5. J. J. Heckman, Skill formation and the economics of investing in disadvantaged children. *Science* **312**, 1900–1902 (2006). Medline doi:10.1126/science.1128898

6. J. J. Heckman, Schools, Skills, and Synapses. *Econ. Inq.* **46**, 289–324 (2008). Medline doi:10.1111/j.1465-7295.2008.00163.x

7. P. Carneiro, J. J. Heckman, in *Inequality in America: What Role for Human Capital Policies?* J. J. Heckman, A. B. Krueger, B. M. Friedman, Eds. (MIT Press, Cambridge, MA, 2003), pp. 77–239.

8. F. Cunha, J. J. Heckman, L. J. Lochner, D. V. Masterov, in *Handbook of the Economics of Education*, E. A. Hanushek, F. Welch, Eds. (North-Holland, Amsterdam, 2006), chap. 12, pp. 697–812.

9. G. J. van den Berg, M. Lindeboom, F. Portrait, Economic conditions early in life and individual mortality. *Am. Econ. Rev.* **96**, 290–302 (2006). doi:10.1257/000282806776157740

10. D. Almond, L. Edlund, H. Li, J. Zhang, Long-term effects of the 1959-1961 China famine: Mainland China and Hong Kong, *Working Paper 13384*, National Bureau of Economic Research (2007).

11. H. Bleakley, Disease and development: Evidence from hookworm eradication in the American South. *Q. J. Econ.* **122**, 73–117 (2007). Medline doi:10.1162/qjec.121.1.73

12. S. L. Maccini, D. Yang, Under the weather: Health, schooling, and economic consequences of early-life rainfall. *Am. Econ. Rev.* **99**, 1006–1026 (2009). doi:10.1257/aer.99.3.1006

13. D. Almond, J. Currie, in *Handbook of Labor Economics*, O. Ashenfelter, D. Card, Eds. (Elsevier, North Holland, 2011), vol. 4B, chap. 15, pp. 1315–1486.

14. P. L. Engle, M. M. Black, J. R. Behrman, M. Cabral de Mello, P. J. Gertler, L. Kapiriri, R. Martorell, M. E. Young; International Child Development Steering Group, Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world. *Lancet* **369**, 229–242 (2007). Medline doi:10.1016/S0140-6736(07)60112-3

15. P. L. Engle, L. C. Fernald, H. Alderman, J. Behrman, C. O'Gara, A. Yousafzai, M. C. de Mello, M. Hidrobo, N. Ulkuer, I. Ertem, S. Iltus; Global Child Development Steering Group, Strategies for reducing inequalities and improving developmental outcomes for

young children in low-income and middle-income countries. *Lancet* **378**, 1339–1353 (2011). Medline doi:10.1016/S0140-6736(11)60889-1

16. J. J. Heckman, Policies to foster human capital. *Res. Econ.* **54**, 3–56 (2000). doi:10.1006/reec.1999.0225

17. S. Grantham-McGregor, Y. B. Cheung, S. Cueto, P. Glewwe, L. Richter, B. Strupp; International Child Development Steering Group, Developmental potential in the first 5 years for children in developing countries. *Lancet* **369**, 60–70 (2007). Medline doi:10.1016/S0140-6736(07)60032-4

18. S. P. Walker, T. D. Wachs, J. M. Gardner, B. Lozoff, G. A. Wasserman, E. Pollitt, J. A. Carter; International Child Development Steering Group, Child development: Risk factors for adverse outcomes in developing countries. *Lancet* **369**, 145–157 (2007). Medline doi:10.1016/S0140-6736(07)60076-2

19. C. Paxson, N. Schady, *J. Hum. Resour.* **42**, 49 (2007).

20. L. C. Fernald, P. Kariger, M. Hidrobo, P. J. Gertler, Socioeconomic gradients in child development in very young children: Evidence from India, Indonesia, Peru, and Senegal. *Proc. Natl. Acad. Sci. U.S.A.* **109** (suppl. 2), 17273–17280 (2012). Medline doi:10.1073/pnas.1121241109

21. J. Heckman, S. H. Moon, R. Pinto, P. Savelyev, A. Yavitz, Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. *Quant. Econom.* **1**, 1–46 (2010). Medline doi:10.3982/QE8

22. J. J. Heckman, S. H. Moon, R. Pinto, P. A. Savelyev, A. Yavitz, The rate of return to the High/Scope Perry Preschool Program. *J. Public Econ.* **94**, 114–128 (2010). Medline doi:10.1016/j.jpubeco.2009.11.001

23. A. J. Reynolds, S.-R. Ou, J. W. Topitzes, Paths of effects of early childhood intervention on educational attainment and delinquency: A confirmatory analysis of the Chicago Child-Parent Centers. *Child Dev.* **75**, 1299–1328 (2004). Medline doi:10.1111/j.1467-8624.2004.00742.x

24. A. J. Reynolds, J. A. Temple, S. R. Ou, D. L. Robertson, J. P. Mersky, J. W. Topitzes, M. D. Niles, Effects of a school-based, early childhood intervention on adult health and well-being: A 19-year follow-up of low-income families. *Arch. Pediatr. Adolesc. Med.* **161**, 730–739 (2007). Medline doi:10.1001/archpedi.161.8.730

25. A. J. Reynolds, J. A. Temple, S.-R. Ou, I. A. Arteaga, B. A. B. White, School-based early childhood education and age-28 well-being: Effects by timing, dosage, and subgroups. *Science* **333**, 360–364 (2011). Medline doi:10.1126/science.1203618

26. F. A. Campbell, C. T. Ramey, E. Pungello, J. Sparling, S. Miller-Johnson, Early childhood education: Young adult outcomes from the Abecedarian Project. *Appl. Dev. Sci.* **6**, 42–57 (2002). doi:10.1207/S1532480XADS0601_05

27. F. A. Campbell, E. P. Pungello, M. Burchinal, K. Kainz, Y. Pan, B. H. Wasik, O. A. Barbarin, J. J. Sparling, C. T. Ramey, Adult outcomes as a function of an early childhood educational program: An Abecedarian Project follow-up. *Dev. Psychol.* **48**, 1033–1043 (2012). Medline doi:10.1037/a0026644

28. F. Campbell, G. Conti, J. J. Heckman, S. H. Moon, R. Pinto, E. Pungello, Y. Pan, Early childhood investments substantially boost adult health. *Science* **343**, 1478–1485 (2014). doi:10.1126/science.1248429 Medline

29. A. Aughinbaugh, Does Head Start yield long-term benefits? *J. Hum. Resour.* **36**, 641 (2001). doi:10.2307/3069637

30. E. Garces, D. Thomas, J. Currie, Longer-term effects of Head Start. *Am. Econ. Rev.* **92**, 999–1012 (2002). doi:10.1257/00028280260344560

31. G. Psacharopoulos, H. A. Patrinos, Returns to investment in education: A further update. *Educ. Econ.* **12**, 111–134 (2004). doi:10.1080/0964529042000239140

32. S. M. Grantham-McGregor, C. A. Powell, S. P. Walker, J. H. Himes, Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: The Jamaican Study. *Lancet* **338**, 1–5 (1991). Medline doi:10.1016/0140-6736(91)90001-6

33. There are, however, experimental studies that show that early-life nutritional interventions also have substantial impacts on earnings (*44*).

34. S. P. Walker, S. M. Chang, M. Vera-Hernández, S. Grantham-McGregor, Early childhood stimulation benefits adult competence and reduces violent behavior. *Pediatrics* **127**, 849–857 (2011). Medline doi:10.1542/peds.2010-2231

35. S. P. Walker, C. A. Powell, S. M. Grantham-McGregor, Dietary intakes and activity levels of stunted and non-stunted children in Kingston, Jamaica. Part 1. Dietary intakes. *Eur. J. Clin. Nutr.* **44**, 527–534 (1990). Medline

36. S. P. Walker, S. M. Chang, C. A. Powell, S. M. Grantham-McGregor, Effects of early childhood psychosocial stimulation and nutritional supplementation on cognition and education in growth-stunted Jamaican children: Prospective cohort study. *Lancet* **366**, 1804–1807 (2005). Medline doi:10.1016/S0140-6736(05)67574-5

37. S. P. Walker, S. M. Chang, M. Vera-Hernández, S. Grantham-McGregor, Early childhood stimulation benefits adult competence and reduces violent behavior. *Pediatrics* **127**, 849–857 (2011). Medline doi:10.1542/peds.2010-2231

38. S. P. Walker, S. M. Grantham-Mcgregor, C. A. Powell, S. M. Chang, Effects of growth restriction in early childhood on growth, IQ, and cognition at age 11 to 12 years and the benefits of nutritional supplementation and psychosocial stimulation. *J. Pediatr.* **137**, 36–41 (2000). Medline doi:10.1067/mpd.2000.106227

39. J. M. Robins, A. Rotnitzky, L. P. Zhao, Estimation of regression coefficients when some regressors are not always observed. *J. Am. Stat. Assoc.* **89**, 846–866 (1994). doi:10.1080/01621459.1994.10476818

40. J. P. Romano, M. Wolf, Exact and approximate stepdown methods for multiple hypothesis testing. *J. Am. Stat. Assoc.* **100**, 94–108 (2005). doi:10.1198/016214504000000539

41. B. M. Caldwell, Descriptive evaluations of child development and of developmental settings. *Pediatrics* **40**, 46–54 (1967). Medline

42. B. M. Caldwell, R. H. Bradley, *HOME Observation for Measurement of the Environment* (University of Arkansas at Little Rock, Little Rock, AR, 1984).

43. S. M. Grantham-McGregor, S. P. Walker, S. M. Chang, C. A. Powell, Effects of early childhood supplementation with and without stimulation on later development in stunted Jamaican children. *Am. J. Clin. Nutr.* **66**, 247–253 (1997). Medline

44. J. Hoddinott, J. A. Maluccio, J. R. Behrman, R. Flores, R. Martorell, Effect of a nutrition intervention during early childhood on economic productivity in Guatemalan adults. *Lancet* **371**, 411–416 (2008). Medline doi:10.1016/S0140-6736(08)60205-6

45. P. V. Hamill, T. A. Drizd, C. L. Johnson, R. B. Reed, A. F. Roche, W. M. Moore, Physical growth: National Center for Health Statistics percentiles. *Am. J. Clin. Nutr.* **32**, 607–629 (1979). Medline

46. I. C. Uzgiris, J. M. Hunt, *Assessment in infancy: Ordinal scales of psychological development.* (University of Illinois Press., Urbana, IL, 1975).

47. S. P. Walker, S. M. Grantham-McGregor, C. A. Powell, J. H. Himes, D. T. Simeon, Morbidity and the growth of stunted and nonstunted children, and the effect of supplementation. *Am. J. Clin. Nutr.* **56**, 504–510 (1992). Medline

48. S. P. Walker, C. A. Powell, S. M. Grantham-McGregor, J. H. Himes, S. M. Chang, Nutritional supplementation, psychosocial stimulation, and growth of stunted children: The Jamaican study. *Am. J. Clin. Nutr.* **54**, 642–648 (1991). Medline

49. J. A. Maluccio, J. Hoddinott, J. R. Behrman, R. Martorell, A. R. Quisumbing, A. D. Stein, The impact of improving nutrition during early childhood on education among Guatemalan adults. *Econ. J.* **119**, 734–763 (2009). doi:10.1111/j.1468-0297.2009.02220.x

50. J. P. Romano, M. Wolf, Stepwise multiple testing as formalized data snooping. *Econometrica* **73**, 1237–1282 (2005). doi:10.1111/j.1468-0262.2005.00615.x